Is automatic or manual transmission better for mileage?

Motor Trend, a US car magazine, is interested in exploring the relationship between the type of transmission (automatic or manual) and fuel consumption/mileage using data they have collected on different US car models in 1973-74. Our exploratory data analysis reveals that, without accounting for other characteristics of automobile design and performance, cars with automatic transmission have significantly worst mileage (-7.24 mpg, 95% confidence interval: -11.28 -3.21 mpg) than cars with manual transmission.

However, our analysis shows that this effect is mainly due to the fact that automatic cars in the seventies were equipped with bigger engines (and thus were also heavier) than manual cars. When accounting for this, the difference in mileage between automatic and manual transmission is in fact statistically insignificant.

This is likely to reflect the technological limitations of automatic transmission design at that time, since these old automatic cars had fewer forward gears and smaller rear axle ratios than manual cars (or even modern automatic cars) and thus required more torque to achieve similar performance/acceleration.

Exploratory data analysis and regression modeling

```
data(mtcars)
head(mtcars, 3)
```

```
##                mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4     21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710    22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
```

The dataset (mtcars) includes fuel consumption in miles per gallon (mpg) and 10 characteristics of automobile design and performance for 32 different US car models in 1973-74. These characteristics are: number of cylinders (cyl), displacement (disp, in cubic inches), horsepower (hp), rear axle ratio (drat), weight (wt, in thousands of pounds), acceleration (qsec, 1/4 mile time in seconds), configuration of cylinders (vs, 0 = v-shaped, 1 = straight/in-line), type of transmission (am, 0 = automatic, 1 = manual), number of forward gears (gear), number of carburetors (carb).

As already mentioned in the executive summary above, cars with automatic transmission have significantly worst mileage than cars with manual transmission (see sample estimates and 95% confidence interval below).

```
t.test(mpg ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  mpg by am
## t = -3.7671, df = 18.332, p-value = 0.001374
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.280194  -3.209684
## sample estimates:
##   mean in group auto mean in group manual
##             17.14737             24.39231
```

```r
fit = lm(mpg ~ am, data = mtcars)
summary(fit)
```

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## ammanual       7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

However, when accounting for the fact that automatic cars in the seventies required bigger engines to achieve similar performance/acceleration (see Appendix), the difference in mileage between automatic and manual transmission is no longer statistically significant.

```r
fit2 = lm(mpg ~ am + disp, data = mtcars)
summary(fit2)
```

```
##
## Call:
## lm(formula = mpg ~ am + disp, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6382 -2.4751 -0.5631  2.2333  6.8386
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 27.848081   1.834071  15.184 2.45e-15 ***
## ammanual     1.833458   1.436100   1.277    0.212
## disp        -0.036851   0.005782  -6.373 5.75e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.218 on 29 degrees of freedom
## Multiple R-squared:  0.7333, Adjusted R-squared:  0.7149
## F-statistic: 39.87 on 2 and 29 DF,  p-value: 4.749e-09
```

```r
anova(fit, fit2)
```

```
## Analysis of Variance Table
```

```
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + disp
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     30 720.90
## 2     29 300.28  1    420.62 40.621 5.748e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
full.model = lm(mpg ~ ., data = mtcars)
model = step(full.model, direction = "backward")
```

```
## Start:  AIC=70.9
## mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
##
##        Df Sum of Sq    RSS    AIC
## - cyl   1    0.0799 147.57 68.915
## - vs    1    0.1601 147.66 68.932
## - carb  1    0.4067 147.90 68.986
## - gear  1    1.3531 148.85 69.190
## - drat  1    1.6270 149.12 69.249
## - disp  1    3.9167 151.41 69.736
## - hp    1    6.8399 154.33 70.348
## - qsec  1    8.8641 156.36 70.765
## <none>              147.49 70.898
## - am    1   10.5467 158.04 71.108
## - wt    1   27.0144 174.51 74.280
##
## Step:  AIC=68.92
## mpg ~ disp + hp + drat + wt + qsec + vs + am + gear + carb
##
##        Df Sum of Sq    RSS    AIC
## - vs    1    0.2685 147.84 66.973
## - carb  1    0.5201 148.09 67.028
## - gear  1    1.8211 149.40 67.308
## - drat  1    1.9826 149.56 67.342
## - disp  1    3.9009 151.47 67.750
## - hp    1    7.3632 154.94 68.473
## <none>              147.57 68.915
## - qsec  1   10.0933 157.67 69.032
## - am    1   11.8359 159.41 69.384
## - wt    1   27.0280 174.60 72.297
##
## Step:  AIC=66.97
## mpg ~ disp + hp + drat + wt + qsec + am + gear + carb
##
##        Df Sum of Sq    RSS    AIC
## - carb  1    0.6855 148.53 65.121
## - gear  1    2.1437 149.99 65.434
## - drat  1    2.2139 150.06 65.449
## - disp  1    3.6467 151.49 65.753
## - hp    1    7.1060 154.95 66.475
## <none>              147.84 66.973
## - am    1   11.5694 159.41 67.384
```

```
## - qsec  1   15.6830 163.53 68.200
## - wt    1   27.3799 175.22 70.410
##
## Step:  AIC=65.12
## mpg ~ disp + hp + drat + wt + qsec + am + gear
##
##        Df Sum of Sq    RSS    AIC
## - gear  1     1.565 150.09 63.457
## - drat  1     1.932 150.46 63.535
## <none>              148.53 65.121
## - disp  1    10.110 158.64 65.229
## - am    1    12.323 160.85 65.672
## - hp    1    14.826 163.35 66.166
## - qsec  1    26.408 174.94 68.358
## - wt    1    69.127 217.66 75.350
##
## Step:  AIC=63.46
## mpg ~ disp + hp + drat + wt + qsec + am
##
##        Df Sum of Sq    RSS    AIC
## - drat  1     3.345 153.44 62.162
## - disp  1     8.545 158.64 63.229
## <none>              150.09 63.457
## - hp    1    13.285 163.38 64.171
## - am    1    20.036 170.13 65.466
## - qsec  1    25.574 175.67 66.491
## - wt    1    67.572 217.66 73.351
##
## Step:  AIC=62.16
## mpg ~ disp + hp + wt + qsec + am
##
##        Df Sum of Sq    RSS    AIC
## - disp  1     6.629 160.07 61.515
## <none>              153.44 62.162
## - hp    1    12.572 166.01 62.682
## - qsec  1    26.470 179.91 65.255
## - am    1    32.198 185.63 66.258
## - wt    1    69.043 222.48 72.051
##
## Step:  AIC=61.52
## mpg ~ hp + wt + qsec + am
##
##        Df Sum of Sq    RSS    AIC
## - hp    1     9.219 169.29 61.307
## <none>              160.07 61.515
## - qsec  1    20.225 180.29 63.323
## - am    1    25.993 186.06 64.331
## - wt    1    78.494 238.56 72.284
##
## Step:  AIC=61.31
## mpg ~ wt + qsec + am
##
##        Df Sum of Sq    RSS    AIC
## <none>              169.29 61.307
```

```
## - am    1    26.178 195.46 63.908
## - qsec  1   109.034 278.32 75.217
## - wt    1   183.347 352.63 82.790
```

```
summary(model)
```

```
##
## Call:
## lm(formula = mpg ~ wt + qsec + am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4811 -1.5555 -0.7257  1.4110  4.6610
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.6178     6.9596   1.382 0.177915
## wt           -3.9165     0.7112  -5.507 6.95e-06 ***
## qsec          1.2259     0.2887   4.247 0.000216 ***
## ammanual      2.9358     1.4109   2.081 0.046716 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.459 on 28 degrees of freedom
## Multiple R-squared:  0.8497, Adjusted R-squared:  0.8336
## F-statistic: 52.75 on 3 and 28 DF,  p-value: 1.21e-11
```

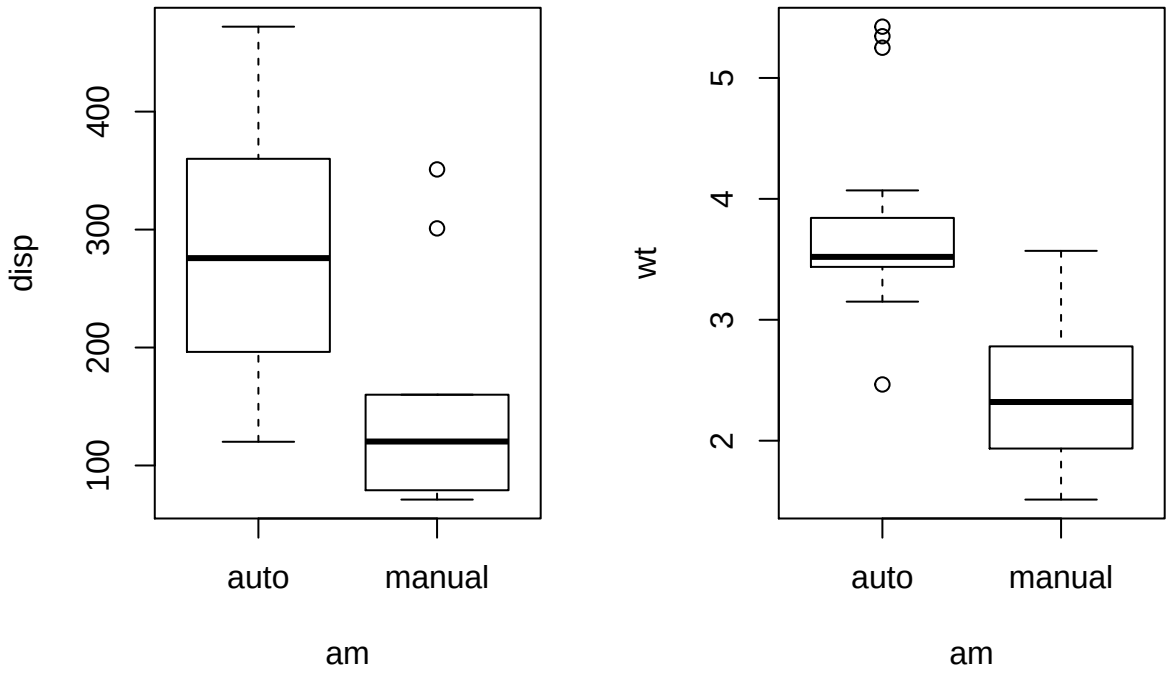
Appendix

Exploratory data analysis

As already mentioned in the executive summary, automatic cars in the seventies were equipped with bigger engines (and thus were also heavier) than manual cars.

```
par(mfrow=c(1,2))

plot(disp ~ am, data = mtcars)
plot(wt ~ am, data = mtcars)
```




```
t.test(disp ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  disp by am
## t = 4.1977, df = 29.258, p-value = 0.00023
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   75.32779 218.36857
## sample estimates:
##   mean in group auto mean in group manual
##             290.3789             143.5308
```

```
t.test(wt ~ am, data = mtcars)
```

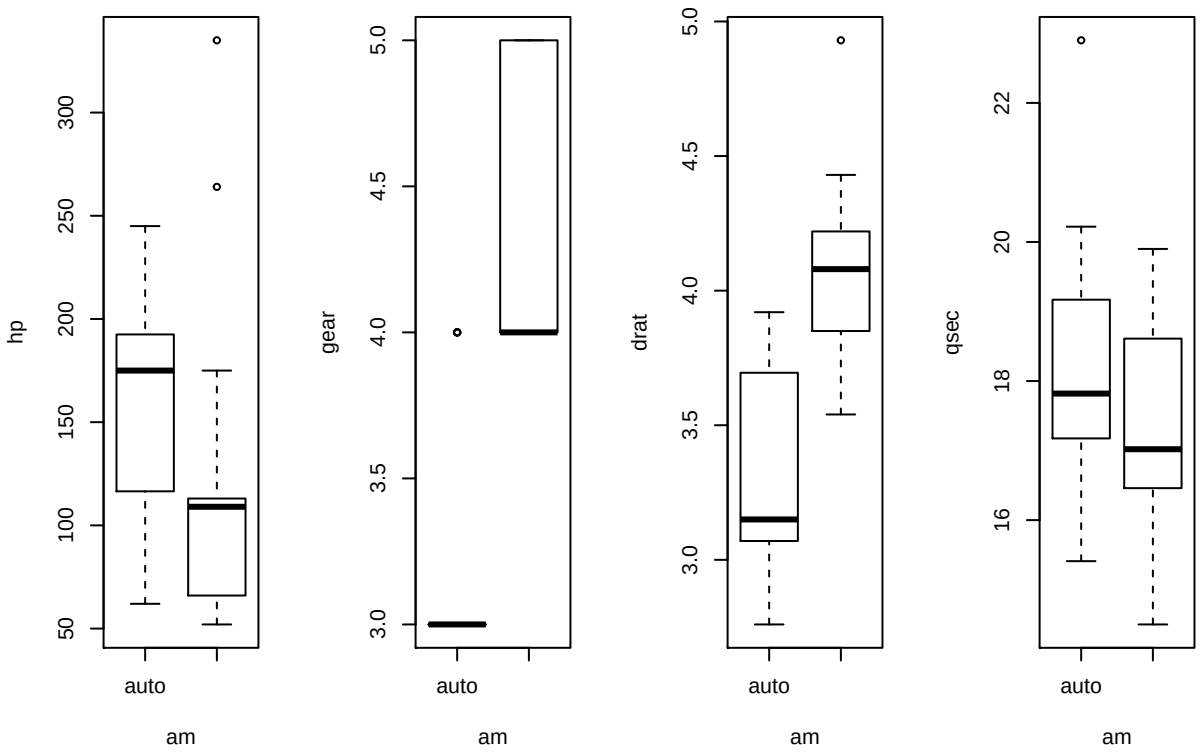
```
##
##  Welch Two Sample t-test
##
## data:  wt by am
## t = 5.4939, df = 29.234, p-value = 6.272e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.8525632 1.8632262
## sample estimates:
##   mean in group auto mean in group manual
##             3.768895             2.411000
```

Given that automatic and manual cars had similar horsepower and performance/acceleration, it appears that they were equipped with bigger engines in order to produce more torque (of two engines with equal horsepower, the larger one will produce more torque). This makes sense because, due to the technological

limitations of automatic transmission design in the seventies, these old automatic cars had fewer forward gears and smaller rear axle ratios than manual cars (or even modern automatic cars) and thus required more torque to achieve similar performance/acceleration.

```
par(mfrow=c(1,4))

plot(hp ~ am, data = mtcars)
plot(gear ~ am, data = mtcars)
plot(drat ~ am, data = mtcars)
plot(qsec ~ am, data = mtcars)
```




```
t.test(hp ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  hp by am
## t = 1.2662, df = 18.715, p-value = 0.221
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -21.87858  88.71259
## sample estimates:
##   mean in group auto mean in group manual
##             160.2632             126.8462
```

```
t.test(gear ~ am, data = mtcars)
```

```
##
```

```
##  Welch Two Sample t-test
##
## data:  gear by am
## t = -6.8995, df = 22.568, p-value = 5.462e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.5264836 -0.8216945
## sample estimates:
##   mean in group auto mean in group manual
##             3.210526             4.384615
```

```
t.test(drat ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  drat by am
## t = -5.6461, df = 27.198, p-value = 5.267e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.0411183 -0.4862501
## sample estimates:
##   mean in group auto mean in group manual
##             3.286316             4.050000
```

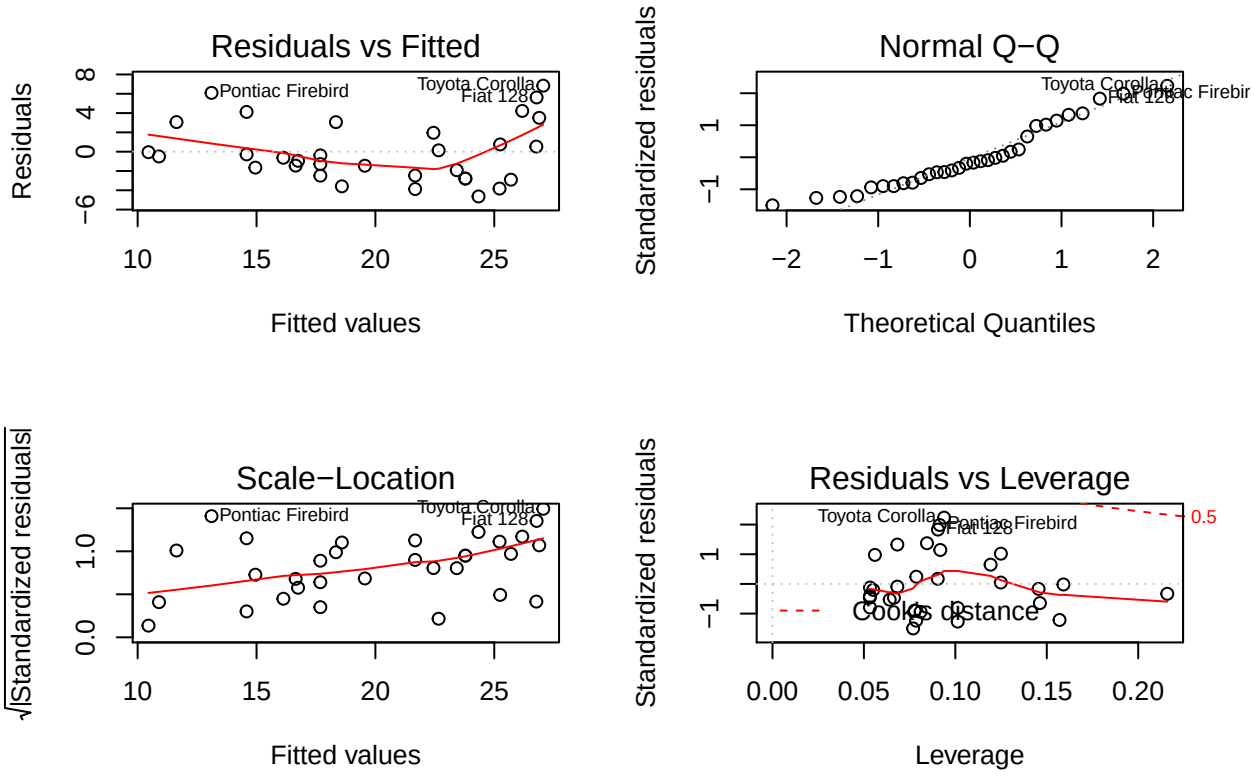
```
t.test(qsec ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  qsec by am
## t = 1.2878, df = 25.534, p-value = 0.2093
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4918522  2.1381679
## sample estimates:
##   mean in group auto mean in group manual
##             18.18316             17.36000
```

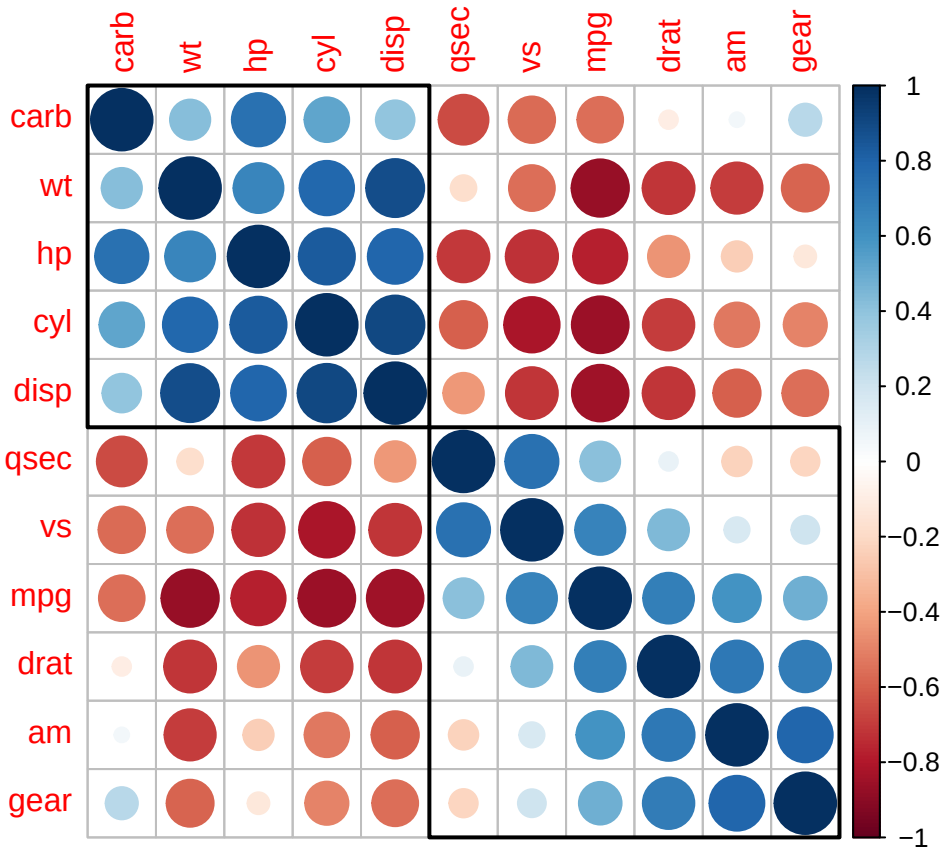
Model diagnostics

```
par(mfrow=c(2,2))

plot(fit2)
```




```
data(mtcars)
library(corrplot)
corrplot(cor(mtcars), order = "hclust", addrect = 2)
```


carb
wt
hp
cyl
disp
qsec
vs
mpg
drat
am
gear
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1